REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
	New Reprint	-

4. TITLE AND SUBTITLE
Improving target detection in visual search through the augmenting multi-sensory cues

5a. CONTRACT NUMBER
W911NF-08-1-0196

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
James Merlo, Joseph E. Mercado, Jan B.F. Van Erp, Peter A. Hancock

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Central Florida
12201 Research Parkway
Suite 501
Orlando, FL 32826 -3246

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
54182-LS.8

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The present experiment tested 60 individuals on a multiple screen, visual target detection task. Using a within-participant design, individuals received no-cue augmentation, an augmenting tactile cue alone, an augmenting auditory cue alone or both of the latter augmentations in combination. Results showed significant and substantive improvements in performance such that successful search speed was facilitated by more than 43%, errors of omission were reduced by 86% and errors of commission were reduced by more than 77% in the

15. SUBJECT TERMS
auditory cueing, tactile cueing, augmented support, target detection, visual search

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Peter Hancock
					19b. TELEPHONE NUMBER 407-823-2310

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Improving target detection in visual search through the augmenting multi-sensory cues

ABSTRACT

The present experiment tested 60 individuals on a multiple screen, visual target detection task. Using a within-participant design, individuals received no-cue augmentation, an augmenting tactile cue alone, an augmenting auditory cue alone or both of the latter augmentations in combination. Results showed significant and substantive improvements in performance such that successful search speed was facilitated by more than 43%, errors of omission were reduced by 86% and errors of commission were reduced by more than 77% in the combinatorial cueing condition compared with the non-cued control. These outcomes were not a trade of performance efficiency for associated mental effort because recorded levels of cognitive workload were also reduced by more than 30% in the multi-cued circumstance compared with the control condition. When the tactile modality was incorporated it led to the highest gain in performance speed, when the auditory modality was incorporated, it led to the best levels of performance accuracy. The combined condition rendered the best of each from of performance increment. Reasons for this outcome pattern are discussed alongside their manifest practical benefits.

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Continuation for Block 13

ARO Report Number 54182.8-LS
Improving target detection in visual search throu ...

Block 13: Supplementary Note

Ergonomics, 2013
Vol. 56, No. 5, 729–738, http://dx.doi.org/10.1080/00140139.2013.771219

Taylor & Francis
Taylor & Francis Group

Improving target detection in visual search through the augmenting multi-sensory cues

Peter A. Hancock[a], Joseph E. Mercado[b]*, James Merlo[b] and Jan B.F. Van Erp[c]

[a]*Department of Psychology, University of Central Florida, Orlando, FL, USA;*
[b]*United States Military Academy, West Point, NY, USA*
[c]*The Netherlands Organization for Applied Scientific Research TNO, Soesterberg, The Netherlands*

(*Received 8 May 2012; final version received 25 January 2013*)

The present experiment tested 60 individuals on a multiple screen, visual target detection task. Using a within-participant design, individuals received no-cue augmentation, an augmenting tactile cue alone, an augmenting auditory cue alone or both of the latter augmentations in combination. Results showed significant and substantive improvements in performance such that successful search speed was facilitated by more than 43%, errors of omission were reduced by 86% and errors of commission were reduced by more than 77% in the combinatorial cueing condition compared with the non-cued control. These outcomes were not a trade of performance efficiency for associated mental effort because recorded levels of cognitive workload were also reduced by more than 30% in the multi-cued circumstance compared with the control condition. When the tactile modality was incorporated it led to the highest gain in performance speed, when the auditory modality was incorporated, it led to the best levels of performance accuracy. The combined condition rendered the best of each from of performance increment. Reasons for this outcome pattern are discussed alongside their manifest practical benefits.

Practitioner Summary: This experiment tested 60 individuals on a multiple screen, visual target detection task. Individuals received no-cue augmentation, tactile cue alone, an augmenting auditory cue alone or both of the latter augmentations in combination. Results showed significant and substantive improvements in the combinatorial cueing condition compared with the non-cued control.

Keywords: auditory cueing; tactile cueing; augmented support; target detection; visual search

Introduction

For human beings, with their extensive emphasis on visual information assimilation (see Sivak 1996), searching environmental displays for critical cues for action is an essential everyday capacity. As such, visual search is a well-researched and progressively more understood response characteristic (Wolfe, Horowitz, and Kenner 2005). Although visual search is often satisfactorily achieved, success is not always assured. Indeed, search failure becomes increasingly more likely when targets to be detected are ambiguous, only marginally above the sensory threshold of observation or physically masked or obscured in some fashion. In addition, visual search becomes increasingly difficult where a large number of targets are presented. Detection capacity also degrades across time when there is an imperative to search for infrequent targets that are embedded in more frequent, non-target distractors. This latter circumstance is a condition that induces the classic vigilance decrement function (Mackworth 1948; Warm 1984; Hancock 2012). In these typical vigilance conditions, visual target detection is also significantly diminished by the presence of accompanying sources of stress (see Hancock and Warm 1989). Failure rate thus increases as the targets to be detected decrease in their sensory and cognitive conspicuity; however, such degradation is generated.

Failure to detect targets has serious consequences in many practical world activities. The results of such failure are, for example, evident in the injuries and casualties of daily car accidents, and being particularly relevant to many modern military endeavours. In contemporary conflicts, for example, it is often the case that enemy combatants are involved in insurgencies that feature munitions now commonly known under the label of, improvised explosive devices (IEDs). These forms of explosive device are truly effective when their location and nature can be hidden or masked from a successful visual search. Of course, these two are just a limited set of exemplars of the tragedies that can follow upon failed visual search. It is the result of these eminently pragmatic and imperative necessities for success in practical visual searches, together with an increasing theoretical interest in sensory cue integration (see Spence 2011), which raises the important questions that have motivated the present work.

Rapid advancements in technology have created new avenues and capacities to detect targets of interest. These opportunities are expressed in different real-world realms ranging from modern vehicles equipped with radar and ultrasonic sensors embedded in collision avoidance systems to displays derived from satellite detection and the contemporary use of

*Corresponding author. Email: jmercado85@gmail.com

unmanned aerial vehicles (Murphy and Bott 1995) as surveillance platforms to support IED detection by ground soldiers. Similar concepts have been introduced to facilitate domestic first responder's reactions to natural disasters as well as dangerous chemical spills and radiation exposures. Of course, they continue to be ever more sophisticated forms of diagnostic procedures in the medical domain (see e.g. Krupinski 2000), these being only a few relevant instances of technological assistance to search capacities. Although the persistent problem in all of these searches is to distinguish targets from non-targets, the proliferation of raw data in all realms constantly threatens to overwhelm the unaided observer. As a result, distinguishing the methods that aid in visual search is practically a very important pursuit. One especially promising avenue is through the provision of multi-modal augmenting cues, which can alert and direct the searcher's visual attention, especially in visual overload situations (Vitense, Jacko, and Emery 2003). Providing such cues or directions to the targets or spatial areas of interest obviously help orient the searcher's attention to the appropriate search region. Orienting cues also provide some enhancement in the general level of observer arousal, an action that itself may serve to facilitate detection.

In respect of detection capacities, Posner, Snyder, and Davidson (1980) distinguished between two different aspects of the attentional system: *orienting* and *detecting*. Orienting denotes where, and in what direction in space, attention should be focused. Detection occurs when there is contact between the attentional system and the signal to be detected (e.g. a crossing pedestrian, the presence of an IED, a combatant, wounded persons, etc.). From these two different aspects of attention, Posner et al. (1980) concluded that the efficiency of target detection is directly affected by orienting and therefore, orienting necessarily either precedes detection or must co-occur in time in order for search to be successful. The preponderance of existing evidence shows that the use of reliable attention cueing that supports *orienting*, albeit even though that cue is somewhat imprecise as to the target's actual location, results in improved response time relative to a no-cued control condition (e.g. Fisher et al. 1989; Fisher and Tan 1989; Hofer, Palen, and Possolo 1993; Merlo and Hancock 2011; Sklar and Sarter 1999; Van Erp et al. 2007). Such basic findings subsequently inform the process of interface design to deal with the practical problems and issues that we now examine.

While advances in technology have made more information available, as well as providing the capability to present that information to the user, the modality of such information presentation is still an interface design choice (Sarter 2006). In many working environments this choice of modality is limited. For example, noisy environments restrict the range of possible auditory information that can be displayed (see Szalma and Hancock 2011, 2012) and, as we have already noted, the vast proliferation of visually presented information often makes the addition of yet another visual display simply impractical. As multi-sensory processors, human operators naturally rely on their differing sensory capacities to integrate the various features of any individual stimulus, or across a spectrum of different stimuli (Philippi, Van Erp, and Werkhoven 2008). They also use these multiple sources to aid them in the initial process of orientation and the subsequent focus of their attention in space and time. When a person directs her or his attention towards a particular location, regardless of the primary modality used in the process of detection, the other modalities are most frequently directed towards that same location also (Ferris and Sarter 2008). These cross-modal, spatial links allow humans to integrate information from several different sensory channels, thus aiding them in constructing an overall representation of space (Driver and Spence 1998; Ernst and Bülthoff 2004). Indeed, more recently, the orientation of attention has been considered as a multi-sensory construction (Spence and Driver 2004) instead of an over-dominantly visual process (Posner, Nissen, and Klein 1976). That the orientation of attention is a multi-sensory construction has also been recently confirmed by neurophysiological investigations (see e.g. Allman and Meredith 2007; Stein and Meredith 1993; Teder-Sälejärvi et al. 2005).

Although the practical advantages of cue augmentation are encouraging, it is still not precisely certain how these advantages are represented in patterns of neurological response. Initially, we might ask whether it is possible to construct an account of the outcome patterns only based on reflections of fundamental properties of each of the peripheral receptor systems. For instance, the known speed advantage of the tactile system may well relate to purely architectural advantage of tactile stimulation over audition. That is, the auditory information has to proceed through an additional step in terms of fundamental anatomical requirements and tactile throughput may, as a consequence simply be faster due to these structural differences. For any associated accuracy effects, one could also postulate a purely structural account also. In typical experiments, tactile stimulation occurs via single tactors and gains no further resolution from any movement of the head or body (i.e. experimental tasks in the tactile situations are often data limited in nature), whereas greater spatial acuity can potentially be gained by head movement during the auditory presentations of the cueing signal; thus, there maybe the resource-based opportunities for greater resolution. However, behavioral data show that even without such improved resolution, the tactile directional cues can be perceived with a high accuracy, probably close to 100% (e.g. see Van Erp 2005). In addition, contrasting peripheral characteristics between receptor systems are not able to explain multi-sensory effects. Taking all this into account makes it implausible that only a simple peripheral-based explanation can account for the multi-modal effects, that is, without involving central information processing structures. Differences between sensory modalities are reflected in the architecture of the involved brain areas (see Spence 2011). Apart from the somatosensory

cortices, we expect that tactile cues will be centrally processed first in the parietal lobe, where the sensory information from the different modalities is integrated. This is particularly so in the case of our mentioned task, which demands the determination of both a spatial sense and dynamic navigation; an evident form of complex visuo-spatial processing. Although multi-sensory in nature, the posterior parietal cortex is often referred to by vision researchers as a part of the dorsal stream of vision (i.e. spatial vision), which arguably plays a major role in the required sensorimotor transformations for visually guided actions, in this case, the direction of visual attention and selection (Goodale and Milner 1992). In contrast, contemporary neurophysiological evidence would indicate that the auditory cue is first processed in the temporal lobe, which referred to as an element of the ventral stream of vision (i.e. perceptual vision). It is from this stream that the brain performs the perceptual identification of objects, thus the efficient pathway for quick identification.

As a consequence of the proceeding observations, tactile cues are thought to 'mediate the required sensorimotor transformations for visually guided actions' (Goodale and Milner 1992). The parietal lobe in the case of the tactile cue is responsible for multi-sensory integration across the bodily senses (e.g. touch to vision), which is a back and forth interaction and is thus a reason that this cue may realise a demonstrated speed advantage. This links the tactile cues to the orienting feature of the attentional system as defined by Posner et al. (1980). In contrast, the auditory cue is in the 'ventral stream of projections from the striate cortex to the infero-temporal cortex playing the major role in the perceptual identification of objects' (Goodale and Milner 1992). This links the auditory cues to the *detecting* feature of the attentional system, which suggests that the auditory cue may possess a superior propensity for identification accuracy. With respect to multi-sensory integration, the human brain constantly integrates sensory information into a holistic view of the world (Ernst and Bülthoff 2004). This integration is automatic for both congruent and incongruent information. However, this integration is not a simple combination of cues across modalities but includes cross weightings of such cues. Models that describe this cue weighting can be summarised by the notion that the most reliable cue has the largest influence in minimising the variance in the final estimate (e.g. Ernst and Banks 2002; Van Erp and Van Veen 2006). In other words, the brain is tuned to seek the optimisation of the best of all sensory facets.

Problematically though, multi-modal stimulation in the real world is not always presented or received in a congruent spatial and temporal manner. This ambiguity may be resolved by over-reliance on the one single dominant system, which in humans is often, but not necessarily always, the visual modality (Hancock 2005, 2010; Werkhoven, Van Erp, and Philippi 2009). However, when there is a strong expectation from past experience that real-world multi-sensory information will be congruent, the benefits should be readily measurable. For example, Glumm, Kehring, and White (2009) conducted a study using U.S. Army personnel and found that the visual cues, spatial tones and haptic cues significantly reduced the amount of time for a M1A1 tank gunner to engage an enemy combatant. The study also showed that the visual and auditory cue times to first shot were equivalent followed by the haptic cues and finally the non-spatial cues. Another study conducted by Van Erp and Van Veen (2004) took into account haptic processes by testing response time in a driving simulator. Navigation directions were given via strictly visual, strictly haptic or multi-modal (a combination of both) avenues. They found that the reaction time was 15% faster when the participant used the multi-modal directions compared with the visual directions alone. Results for the haptic only condition lay between multi-modal and visual conditions. These findings suggest that response time is faster when using haptic cues than while using visual cues, with the combination of the two being even faster. In addition, it put forward mixed results as to whether tactile or auditory cues are better in assisting response speed in a visual search task. Our research intends to fill this gap along with a concomitant and detailed analysis of associated response accuracy.

Although most studies have provided evidence that information presented in multiple modalities is effective, Santangelo and Spence (2007, 1312) have stated that 'combined auditory and visual cues appear to be somewhat less effective in capturing people's attention' than auditory and visual cues by themselves. Therefore, although benefits of multi-modal cueing have often been advocated, the purported benefits are not without criticism. Furthermore, the manner in which, and the degree to which, auditory and tactile cues facilitate complex visual search have yet to be fully explored, explained and exploited (see also Oron-Gilad et al. 2007). Although a number of studies have examined the nature of redundancy across different modality sources (e.g. Calhoun et al. 2004), there have been relatively few studies that have examined redundancy while cues from multiple modalities are presented in a coincident manner (for an exception see Oskarsson, Eriksson, and Carlander 2012). The apparent drawback from such multiple presentations would seem to be the costs associated with a degree of sensory confusion. This would be especially true if the various respective sensory channels communicated information, which was inconsistent either in the spatial or temporal domains. However, in our previous work, we have been encouraged by the improvements encountered in bi-modal forms of presentation (Merlo, Duley, and Hancock 2010) and thus the present extension into the exploration of potential advantages of the tri-modal form of support (see also Oskarsson et al. 2012).

Therefore, the primary purpose of the present work was to examine such cross-modal cueing effects in circumstances that used direct, meaningful and real-world signals. In these more applied settings, the cross-modal advantage of the integration of visual, tactile and auditory information, if confirmed, could improve significantly on single modality communication (Prewett et al. 2012). Such an advantage would be especially evident when any one sensory channel is overloaded or otherwise degraded by local masking or degrading circumstances. For example, the risk of visual overload in

car driving perpetuates a significant threat to traffic safety (De Vries, Van Erp, and Kiefer 2009; Hogema et al. 2009), and in extreme operational conditions, such as combat or firefighting, the capacity to create and retain some form of redundancy gain is not merely useful, but it may prove critical to the survival (Merlo, Szalma, and Hancock 2007). This pursuit of an increased communication capacity is important because missed or misinterpreted signals or messages in such situations often have catastrophic consequences (Reason 2008). From the forgoing, we hypothesised here that there will be an overall benefit in performance for augmented cueing and further that the elements of that benefit in terms of response speed, and response accuracy will be differentially affected by the specific mode (tactile vs. auditory) through which the augmented cue is delivered. In addition, we hypothesised that workload would be reduced with each form of augmenting cue in accordance with their influence on performance efficiency, thus providing a direct and associative effect in this experimental circumstance (see Hancock 1996).

Experimental method

Experimental participants

Sixty cadets (10 females and 50 males) who were college freshers enrolled in a general psychology class at the United States Military Academy (USMA) at West Point, New York participated in the present experimental procedure. These participants aged from 18 to 22 years and had little or no previous experience in monitoring multiple visual information display systems and were thus considered naïve or novice performers. Participants received the extra credit points that counted towards their overall general psychology class grade, participated voluntarily and were treated under the ethical standards rubric of the American Psychological Association. The experiment was approved by the USMA Human Subjects Use Committee and by the Human Subject Committee of the University of Central Florida.

Experimental apparatus

The apparatus used in this experiment included three Dell LCD video monitors, three Altec Lansing FX 4021 speakers and a wearable EAI tactor belt with tactile actuators embedded. These facilities were controlled by a purpose-created, LabView-based software computer program that synchronised the respective displays and recorded response times and accuracy for each participant in identifying the target visual stimuli. The centre screen was directly in front of the participant and approximately 16 inches away from their eyes. Two screens were presented adjacent to the centre screen, one to the left and one to the right. Each of the three visual displays presented different visual search tasks. All participants were unaware of the reliability of the cueing automation, which in this experiment was set to 100%. The reason for setting reliability at this level is that in most practical circumstances even 99% reliability is frequently considered insufficient because it can result in catastrophic errors. Screen one (to the left) always displayed the text messaging ‘chat room’. The participant’s task was to monitor this display for the occurrence of all text messages from ‘Bulldog 6’, which were embedded among the other distractor text messages presented during each trial. Whenever such a message occurred the individual was to click the ‘Acknowledge’ button. Screen two (in the centre) always displayed the view from a driver’s perspective of looking out the front windshield of a vehicle while driving along a specific route. The task here was to ‘Acknowledge’ the occurrence of a specific route marker given to them before each trial, which appeared at sporadic intervals. Once the participant was given a specific route marker, all other routes in the trial then represented non-targets. Screen three (to the right) always displayed a blue force tracker system, that is a top-down map view displaying symbols for friendly and hostile entities. An ‘Acknowledge’ response was required each time any symbol dropped onto the map.

The vibrotactile actuators (tactors) in our tactile communication system were model C2, manufactured by Engineering Acoustics, Inc. These tactors presented 250-Hz sinusoidal vibrations onto the skin through a contactor (diameter 7 mm, with a 1-mm gap separating it from the tactor aluminium housing). Eight tactors were embedded in a belt made of elastic and high-quality cloth similar to the material used by professional cyclist. When stretched around the torso and fastened, the wearer has one actuator over the umbilicus and one centred over his or her spine in the back, whereas the rest are equally distributed around the front. The torso has been found to be a stable and effective reception area and is particularly suited for cueing direction (Redden et al. 2007). In this experiment only three tactors were used, and these were located on the umbilicus, on the left, and on the right side of the torso. Tactile cues were single bursts of 250 Hz lasting 500 ms that occurred in one of the three corresponding spots on the abdomen as the visual screen that was being cued, that is left, right and centre. An Altec Lansing FX 4021 sound system with three speakers was used. Audio messages were a single 900-Hz auditory cue from one speaker at 50 dB lasting 500 ms that could emanate from beneath each of the three corresponding LCD screens. Although the auditory cue matched the target screen with respect to location and direction both, the tactile cue was directional only but not specifically matched with the visual target location. That is, the used tactile actuators were located on the torso, and the actuators linked to the left and right screen were at $\pm 90°$ angles and not at $\pm 22.5°$as the visual

Figure 1. Experimental task and environment. Shown are three monitors, each with a speaker mounted below, keyboard and mouse. Also shown are the tactor belt with battery pack and reference sheet illustrating visual targets that appear on screens two (route markers) and three (blue force tracker symbols) (Colour online).

displays. The combined condition represented the presentation of both the auditory and tactile cue together. All cues were presented simultaneously with the stimulus and were amply above threshold to account for saliency concerns. Figure 1 illustrates the experimental task and environment.

Experimental design

The independent variable in this experiment was type of cueing (i.e. no cueing, tactile cueing alone, auditory cueing alone, tactile and auditory cueing together) to support visual search for target identification across the three respective screens. All 60 participants completed one scenario in each of the four cueing conditions. The dependent variables were the response time, accuracy rate, type of task, location of task, experienced cognitive workload as assessed by the NASA-Task Load Index (TLX; Hart and Staveland 1988) and perceived cue utility as indicated by the participant via a response questionnaire.

There were 15 targets presented in only one scenario, and they were divided such that five targets appeared on each of the three respective screens. Stimuli were presented at the irregular intervals throughout each individual participant's series of trials so that for any single participant there was no identifiable temporal pattern. In the cued conditions, both the tactile and auditory cues were presented simultaneously with the stimuli. The issue of task difficulty and the potential for asymmetric transfer effects were addressed in the following manner. First, the scenarios had been previously evaluated to match for level of difficulty (see Merlo and Hancock 2011) and were counterbalanced across individual participant presentation. Although the issue of potential transfer cannot be solved algorithmically, there are strategic ways of reducing its impact on outcome results (Poulton 1982). As a result, in this experiment, our participants were divided into groups of 15 who undertook the sequence of different scenarios in differing test orders. Each of the four groups was assigned a different sequence of scenario by cueing conditions, and these are specified in Table 1.

Experimental procedure

The experiment was conducted in a controlled, laboratory environment free of competing noise or vibration. Before beginning any of the tasks, the participant was given a short briefing to explain their role in monitoring three video screens and signed the informed consent materials. They were shown precisely how to physically respond by clicking the 'Acknowledge' button, via

Table 1. Participant block group. There were 15 participants per group (total $N = 60$).

Group 1	Scenario 1 (−)	Scenario 2 (+)	Scenario 3 (+) (*)	Scenario 4 (*)
Group 2	Scenario 2 (+) (*)	Scenario 1 (*)	Scenario 4 (−)	Scenario 3 (+)
Group 3	Scenario 3 (*)	Scenario 4 (+) (*)	Scenario 1 (+)	Scenario 2 (−)
Group 4	Scenario 4 (+)	Scenario 3 (−)	Scenario 2 (*)	Scenario 1 (+) (*)

Note: (−), No tactor belt or auditory cueing; (+), tactor belt; (*), auditory cueing; (+) (*), tactor belt and auditory cueing together.

a mouse click on the respective screen that was displaying each pre-specified target. The participant was informed to respond as quickly and as accurately as possible. The participant was also shown representative examples of each of the targets such that they could properly identify each target cue before responding. Finally, the participants were informed as to the nature of each augmenting cue and how it related to the three visual display screens in front of them. Participants were not made aware of any potential failure rate of any augmenting cues. However, in the present experiment, for the purposes of ecological validity, no cue provided incorrect information. On completing the instruction set, participants began the experiment itself. Each individual test scenario lasted approximately 5 min. In general, this task, which resembled actual operation conditions, can be considered as imposing a medium level of demand on the observing individual. Once a participant had completed each scenario, they filled out the NASA-TLX specific to that particular scenario and then moved to the next scenario. A short break was taken after the first two scenarios after which the third and fourth scenarios were completed. After the participants complete testing all four scenarios, they ordered the four cueing conditions regarding their utility for the visual task tested. The participant was then debriefed, thanked and allowed to depart the experiment.

Experimental results

Objective performance

In the present experiment, the objective performance capacity was assessed through three primary dependent measures. These were: (1) response time (defined as the latency between the onset of the stimulus and the subsequent depression of the response button), (2) response omissions (misses) and (3) false alarms (incorrect identifications when signals had not appeared). A one-way multivariate analysis of variance revealed a significant multivariate main effect of cueing type, Wilks' $\lambda = 0.666$, $F(9, 236) = 11.50$, $p < 0.001$. In respect of response time, there was a significant influence of cueing, $F(3, 236) = 36.68$, $p < 0.001$. Here, we found that response time in the non-cued condition was significantly higher, and thus worse, than the response time in any of the other three conditions (i.e. no cue = 3.41 s, tactile cue = 1.94 s, auditory cue = 2.12 s and combined cue = 1.93 s). There proved to be no significant difference between the response times for any of the latter respective cued conditions. However, the mean response time in the combined condition appears to be very close to that for the tactile only cued response.

Analysis of the misses showed a significant effect of cue format, $F(3, 236) = 3.91$, $p = 0.009$. (no cue = 2.33, tactile cue = 1.67, auditory cue = 0.55 and combined cue = 0.33). *Post hoc* comparisons of these miss rates using Tukey's procedure, distinguished significant differences between the no-cue condition and both the auditory cue condition and the combined cue condition. No other pairwise comparisons reached such a significant level of distinction. Thus, we can confirm that the fastest response times were accompanied by the lowest rates of omission (miss) errors and vice versa. For the false alarm rate, we saw a significant effect of condition, $F(3, 236) = 3.27$, $p = 0.022$. Here, we again see that the highest number of false alarms was in the no-cueing condition = 3.0, followed by the tactile condition = 1.56, the auditory condition = 0.78 and the combined cueing condition = 0.67. *Post hoc* analyses using Tukey's procedure showed that there were significant differences between the no-cue condition and the combined cueing condition as well as the no-cue condition and the auditory cue alone. No other pairwise comparisons reached significant levels of difference. The overall outcomes for reflections of response accuracy exhibit no evidence of a speed-accuracy trade-off.

In addition to these foregoing evaluations, we also examined whether the location and type of task differentially affected response capacity. In the present experiment, task type and the physical spatial location of the task were necessarily concatenated; however, we report such effects here in terms of task type. This analysis showed that there was a significant effect in response time for different task types, $F(2, 478) = 27.346$, $p < 0.001$. *Post hoc* analysis of this outcome served to confirm that there was a significant difference between all three tasks such that the response time for the text task (mean = 2.02 s) was significantly faster than that for the driving task (mean = 2.28 s), which in its turn was significantly faster than that for the blue force tracker task (mean = 2.70 s). This outcome was somewhat counterintuitive because the a priori expectation was that the difficulty of the text task was the greatest of the three. We might speculate that these differences reflect a global attention strategy whereby the participant paid their greatest attention to the task of greatest perceived difficulty (as judged by the experimenters and participants in previous experiments using this search configuration). However, this would be to invoke a mediational 'explanation' when no specific test of that capacity has here been undertaken, and also we should not forget to emphasise that such task difficulty was also embedded in the right, left and centre location of each respective task. Suffice it to say that such effects impel us to further empirical exploration.

Subjective ratings

In the present experiment, subjective mental workload was assessed using scores derived from the NASA-TLX (Hart and Staveland 1988) These values exhibited a significant main effect of cueing condition, i.e. $F(3, 236) = 12.64$, $p < 0.001$. For

the overall TLX score, we see a familiar outcome pattern i.e. no cue = 35.63, tactile cue = 24.62, auditory cue = 23.09 and the combined cue condition = 21.34. *Post hoc* analyses of these scores using Tukey's procedure distinguished the workload score in the no-cue condition as being significantly higher than all of the cued conditions. No other pairwise comparison reached significant levels of difference. Such figures indicate a 40.1% reduction in overall global workload score as a result of the most advantageous cue combination. As may also be observed, the outcome for the global workload scores tended very much to follow that specifically for response error in this experiment. We also conducted a direct evaluation of the participants' responses on each the six subscales, which compose the TLX overall score. The six subscales are mental demand, temporal demand, effort, frustration, physical demand and own performance. Respectively, with regards to these six scales, we observed significant effects for four, i.e. mental demand, $F(3, 236) = 16.98$, $p < 0.001$; temporal demand, $F(3, 236) = 7.12$, $p < 0.001$; effort, $F(3, 236) = 12.92$, $p < 0.001$; frustration, $F(3, 236) = 10.39$, $p < 0.001$. Neither physical demand nor own performance showed any significant variation in respect of cue condition. Pairwise comparisons confirmed the pattern that has previously been described above for the overall workload perpetuated into each subscale, that is, the no-cueing condition proved to have significantly higher levels of workload on each subscale as compared with each of the other cued conditions, which did not differ significantly among themselves. We can thus conclude from these overall findings that the superior performance, which was evident in the objective forms of assessment, is not achieved simply by a trade of increased performance capacity for increasing workload but is, in actuality, a case of performance-workload association (and see Hancock 1996).

Finally, we assessed a user preference. To determine this, we asked the participants to rank order their user experience in respect of the four different cueing conditions. The results of this ranking showed that the least preferred condition on a 1–4 (most preferred to least preferred) scale was the no-cue condition (3.55); the next least preferred condition was the tactile only cue (2.62), and this was very close in rank to the next least preferred condition, which was the auditory only (2.45). This left the combined cueing condition, using both the auditory and tactile input as by far the most preferred user condition (1.38). These results confirm that along with strong percentage gains in objective performance and concomitantly decreased levels of associated mental workload, users also preferred the joint cueing condition above any other. Thus, performance and workload advantages were not at the expense of the user acceptance. Overall, these results are highly supportive of combinatorial cueing for increased performance capacity in the visual search and detection arena in which the present procedure was set. The degree to which such advantages extend to other realms of performance await further evaluation; however, we suspect such advantages do persist across a wide range of real-world, operational tasks.

Discussion

What we see in the objective performance pattern is a clearly demonstrable advantage of cue augmentation. Compared with the no augmentation condition, there are significant and meaningful overall performance gains (see also Fisher et al. 1989; Merlo and Hancock 2011; Tindall-Ford, Chandler, and Sweller 1997). We see a response speed benefit of cueing which is largely independent of the precise form of the cue. The mean response time in the combined condition is very close to that for the tactile-only cued responses. This outcome pattern may suggest a form of horse-race model in which reaction occurs in response to stimulation from the fastest cued sensory channel, although it must be reiterated that no formal *post hoc* differences were evident across any of the respective augmented cueing conditions. At its maximum across the all-cues condition, such cueing improved response speed by 43%, and it reduces missed signals by 86% and false alarms by 77%. Each represent important performance gains and thus have very practical impacts on the design of augmented alerting and warning systems. Also each of the performance measures shows that the advantage in the combined cueing condition appears to represent the best gain possible from cue presence in either the tactile or audio cue condition alone. Thus, speed of response is most facilitated by the tactile cue and performance speed in the combined condition is equal to this greatest tactile response speed value (*cf.* Santangelo and Spence 2007). For the accuracy measures, the combined condition proves to be very close to that with the auditory alarm cue alone, see Figure 2. Thus, results confirm a strong advantage in the objective performance that appears to be modality specific for each dimension of that performance.

In this experimental paradigm, the cueing served to move attention laterally across the three-screen display space. Each of tactile and auditory cues in our work only presented information to shift visual attention from the body centre laterally to 22.5° left and right. However, this was still enough to obtain substantial performance gains. The way in which these benefits are derived and the way in which each modality of augmentation contributes to the improvements in speed and in accuracy, combined with the concomitant effects on mental workload ratings represent a distinct and new pattern of findings.

As noted, response speed was facilitated by all forms of cue in an undifferentiated manner. Initially, this might appear to be dissimilar to findings from our own previous research (Merlo et al. 2010; Merlo and Hancock 2011) and that of other groups who have provided definite demonstrations of the speed advantage of tactile cueing (see e.g. Ferris and Sarter 2008;

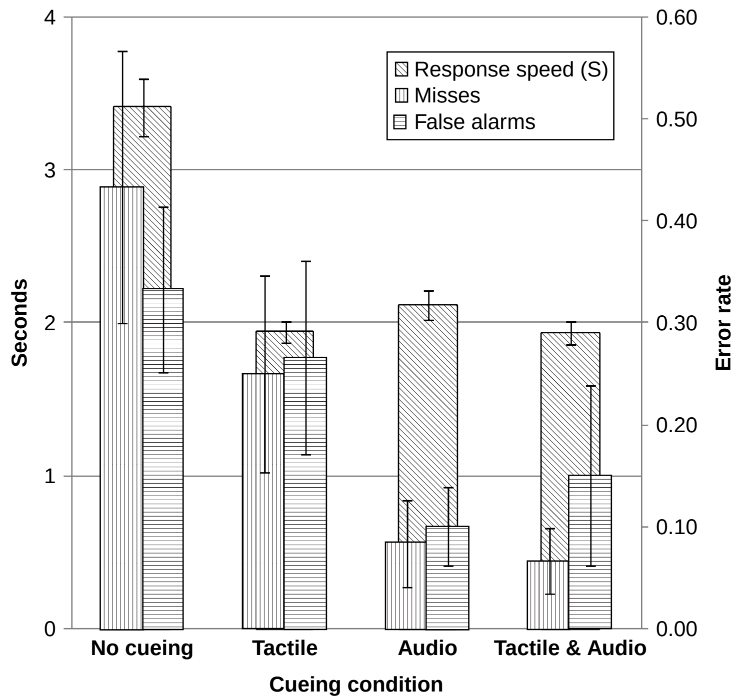

Figure 2. Objective performance measures showing improvement in speed of response and rates of signal misses and false alarms with the introduction of tactile, auditory and combined tactile and auditory cueing.

Jones and Sarter 2008; Mohebbi, Gray, and Tan 2009; Van Erp et al. 2007). The fact that the current data cannot unequivocally substantiate the superior response speed with tactile cueing over auditory cueing may be because of our particular implementation of the different cues in this experiment. Although the auditory cue matched the target screen with respect to both location and direction, our current tactile cue was directional only but not specifically (i.e. spatially homeomorphically) matched to visual target location. That is, the used tactile actuators were located on the torso, and the actuators linked to the left and right screen were at $\pm 90°$ angles and not at $\pm 22.5°$ as were the visual displays themselves. However, despite this particular difference in spatial cue mapping quality, we still see that the tactile cueing is at least as fast as the auditory cueing. With improved spatial resolution of the tactile cue, the speed advantage noted in our own work and that of others may then be restored.

Although the majority of cueing research is restricted to effects on response times, here we explicitly examined the potential trade-off with response accuracy and mental workload in this work. Our results showed that the auditory cues provided the strongest and most consistent improvements in response accuracy and that response accuracy was not influenced by the addition of the tactile cue. This distinction was confirmed in the statistical differentiation of these respective conditions by *post hoc* analysis. Most significantly, in both theoretical and practical terms, each of these respective advantages in speed and accuracy are captured to the greatest degree in the combined cueing condition. Thus, when both augmented cues are added together then the postulated advantage that is derived from the tactile cue in response speed is preserved as is the advantage for response accuracy experienced from auditory augmentation. With respect to the objective reflections of performance, we can thus assert that there is no speed-accuracy trade-off present in the augmenting cue implementation. This means that the practical gains observed are not a result of a strategy change by participating individuals but are actual objective gains. In addition, the recorded mental workload ratings show a similar pattern to the performance measures with the highest load ratings in the no-cued condition. This indicates that the performance gains of cueing do not come at the cost of increased mental workload; a phenomenon not previously establish despite the cues do serve to add yet more information to the interface and therefore ostensibly increase objective task demand. Finally, the descriptive survey results demonstrate that along with reducing the effort associated with complex search, the individuals also preferred this combinatorial cue configuration. Based on these results, we conclude that multi-sensory cues are highly effective in providing support for complex visual search tasks in improving speed and accuracy of responses, in reducing mental workload and in increasing user acceptance.

Conclusions

Multi-sensory audio/tactile cueing improves visual search in terms of speed and accuracy and reduces the amount of mental workload required. When vision is disrupted by glare, sandstorms, night-time conditions and the like, augmented cues in the other senses can make up for what can sometimes be critical operational shortfalls (see Jones and Sarter 2008). Indeed, these benefits of augmented cueing are most likely to emerge in the face of the most disruptive of environments such as emergency rescue or special operations, which are characterised by high-stress imposition (see Hancock and Warm 1989). Such cueing can also help in the distribution of excessive task demand, frequently represented in many modern work systems by the visual overload. This opportunity is supported here by the subjective workload findings, which imply that the reduction experienced under cueing conditions results in the liberation of additional effort that can be used on other necessities. The practical benefits do not apply only, or even primarily, in the tested conditions. The combinatorial modality benefit that was realised may be absolutely essential when each of the discrete sensory channels is masked for various reasons in the real world. Thus, redundant forms of sensory cue enable the individual to feel when it is too noisy to hear and so forth (Szalma and Hancock 2011). This notion of redundancy gain is a stalwart principle that has been used by design professionals across the years. How to maximise this gain through variation of the intensity, saliency and specific informational content of each form of cue augmentation awaits further exploration, explication and exploitation.

Acknowledgements

We would like to thank Cadets from the Engineering Psychology Program at the USMA for their diligence in creating the stimuli for the experiment and helping to collect the data for the present experiment. Our thanks are also due to Dr Aaron Duley for his LabView programming, which enabled the experimental protocol. The work was supported by the government contract number W911NF-08-1-0196, *Adaptation of Physiological and Cognitive Workload via Interactive Multi-Modal Displays* from the Army Research Office, P.A. Hancock, *Principal Investigator*. Dr Elmar Schmeisser was the Technical Monitor for the grant. The views expressed in this work are those of the authors and do not necessarily reflect official Army policy.

References

Allman, B. L., and M. A. Meredith. 2007. “Multisensory Processing in ‘Unimodal’ Neurons: Cross-Modal Subthreshold Auditory Effects in Cat Extrastriate Visual Cortex.” *Journal of Neurophysiology* 98 (1): 545–549.

Calhoun, G. L., J. V. Fontejon, M. H. Draper, H. A. Ruff, and B. J. Guilfoos. 2004. “Tactile Versus Aural Redundant Alert Cues for UAV Control Applications.” *Proceedings of the Human Factors and Ergonomics Society* 48 (1): 137–141.

De Vries, S. C., J. B. F. Van Erp, and R. Kiefer. 2009. “Direction Coding Using a Tactile Chair.” *Applied Ergonomics* 40: 477–484.

Driver, J., and C. Spence. 1998. “Attention and the Cross-Modal Construction of Space.” *Trends in Cognitive Science* 2: 254–262.

Ernst, M. O., and M. S. Banks. 2002. “Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion.” *Nature* 415: 429–433.

Ernst, M. O., and H. H. Bülthoff. 2004. “Merging the Senses into a Robust Percept.” *Trends in Cognitive Science* 8: 162–169.

Ferris, T. K., and N. B. Sarter. 2008. “Crossmodal Links Between Vision, Audition, and Touch in Complex Environments.” *Human Factors* 50 (1): 17–26.

Fisher, D. L., B. G. Coury, T. O. Tengs, and S. A. Duffy. 1989. “Minimizing the Time to Search Visual Displays: The Role of Highlighting.” *Human Factors* 31 (2): 167–182.

Fisher, D. L., and K. C. Tan. 1989. “Visual Displays: The Highlighting Paradox.” *Human Factors* 31: 17–30.

Glumm, M. M., K. L. Kehring, and T. L. White. 2009. “Effects of Unimodal and Multi-Modal Cues About Threat Locations on Target Acquisition and Workload.” *Military Psychology* 21 (4): 497–512.

Goodale, M., and A. Milner. 1992. “Separate Visual Pathways for Perception and Action.” *Trends in Neuroscience* 15 (1): 20–25.

Hancock, P. A. 1996. “Effect of Control Order, Augmented Feedback, Input Device and Practice on Tracking Performance and Perceived Workload.” *Ergonomics* 39: 1146–1162.

Hancock, P. A. 2005. “Time and the Privileged Observer.” *Kronoscope* 5 (2): 176–191.

Hancock, P. A. 2010. “The Battle for Time in the Brain.” In *Time, Limits and Constraints: The Study of Time XIII*, edited by J. A. Parker, P. A. Harris, and C. Steineck, 65–87. Leiden: Brill.

Hancock, P. A. 2012. “In Search of Vigilance; The Problems of Iatrogenically Created Psychological Phenomenon.” *American Psychologist*, in press.

Hancock, P. A., and J. S. Warm. 1989. “A Dynamic Model of Stress and Sustained Attention.” *Human Factors* 31: 519–537.

Hart, S. G., and L. E. Staveland. 1988. “Development of a Multi-Dimensional Workload Rating Scale: Results of Empirical and Theoretical Research.” In *Human Mental Workload*, edited by P. A. Hancock, and N. Meshkati, 139–183. Amsterdam: Elsevier.

Hofer, E. F., L. A. Palen, and A. Possolo. 1993. “Flight Deck Information Management: An Experimental Study of Functional Integration of Approach Data.” In *Proceedings of the 7th International Symposium on Aviation Psychology*. Columbus, OH: Department of Aviation, Ohio State University.

Hogema, J. H., S. C. De Vries, J. B. F. Van Erp, and R. J. Kiefer. 2009. “A Tactile Chair for Direction Coding in Car Driving: Field Evaluation.” *IEEE Transactions on Haptics* 2 (4): 181–188.

Jones, L. A., and N. Sarter. 2008. “Tactile Displays: Guidance for Their Design and Application.” *Human Factors* 50 (1): 90–111.

Krupinski, E. A. 2000. “The Importance of Perception Research in Medical Imaging.” *Radiation Medicine* 18 (6): 329–334.

Mackworth, N. H. 1948. “The Breakdown of Vigilance During Prolonged Visual Search.” *Quarterly Journal of Experimental Psychology* 1: 6–21.

Merlo, J. L., A. R. Duley, and P. A. Hancock. 2010. “Cross-Modal Congruency Benefits for Combined Tactile and Visual Signaling.” *American Journal of Psychology* 123 (4): 413–424.

Merlo, J. L., and P. A. Hancock. 2011. “Quantification of Tactile Cueing for Enhanced Target Search Capacity.” *Military Psychology* 23 (2): 137–153.

Merlo, J., M. Szalma, and P. A. Hancock. 2007. “Stress and Performance: Some Experiences from Iraq.” In *Performance Under Stress*, edited by P. A. Hancock, and J. L. Szalma. Aldershot: Ashgate.

Mohebbi, R., R. Gray, and H. Z. Tan. 2009. “Driver Reaction Time to Tactile and Auditory Rear-End Collision Warnings While Talking on a Cell Phone.” *Human Factors* 51 (1): 102–110.

Murphy, D., and J. Bott. 1995. “On the Lookout: The Air Mobile Ground Security and Surveillance System (AMGSSS) Has Arrived.” *Unmanned Systems* 13 (4): 22–27.

Oron-Gilad, T., J. L. Downs, R. D. Gilson, and P. A. Hancock. 2007. “Vibrotactile Guidance Cues for Target Acquisition.” *IEEE Transactions on Systems, Man, and Cybernetics* 37 (5): 993–1004.

Oskarsson, P-A., L. Eriksson, and O. Carlander. 2012. “Enhanced Perception and Performance by Multimodal Threat Cueing in Simulated Combat Vehicle.” *Human Factors* 54 (1): 122–137.

Philippi, T. G., J. B. F. Van Erp, and P. J. Werkhoven. 2008. “Multisensory Temporal Numerosity Judgment.” *Brain Research* 1242 (C): 116–125.

Posner, M. I., M. J. Nissen, and R. Klein. 1976. “Visual Dominance: An Information-Processing Account of its Origins and Significance.” *Psychological Review* 83: 157–170.

Posner, M. I., C. R. R. Snyder, and B. J. Davidson. 1980. “Attention and the Detection of Signals.” *Journal of Experimental Psychology: General* 109 (2): 160–174.

Poulton, E. C. 1982. “Influential Companions: Effects of One Strategy on Another in the Within-Subjects Designs of Cognitive Psychology.” *Psychological Bulletin* 91 (3): 673.

Prewett, M. S., L. R. Elliott, A. G. Walvoord, and M. D. Coovert. 2012. “A Meta-Analysis of Vibrotactile and Visual Information Displays for Improving Task Performance.” *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 42 (1): 123–132.

Reason, J. 2008. *The Human Contribution*. Chichester: Ashgate.

Redden, E. S., C. B. Carstens, D. D. Turner, J. C. Brill, S. Stafford, and P. I. Terrence. 2007. “Placement, Fit, and Comparison of Two Types of Tactile Displays.” In *Remote Tactile Displays for Future Soldiers*, edited by R. D. Gilson, E. S. Redden, and L. R. Elliott. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Santangelo, V., and C. Spence. 2007. “Multisensory Cues Capture Spatial Attention Regardless of Perceptual Load.” *Journal of Experimental Psychology: Human Perception and Performance* 33 (6): 1311–1321.

Sarter, N. 2006. “Multimodal Information Presentation: Design Guidance and Research Challenges.” *International Journal of Industrial Ergonomics* 36: 439–445.

Sivak, M. 1996. “The Information That Drivers Use: Is It Indeed 90% Visual?” *Perception* 25 (9): 1081–1089.

Sklar, A. E., and N. B. Sarter. 1999. “Good Vibrations: Tactile Feedback in Support of Attention Allocation and Human-Automation Coordination in Event-Driven Domains.” *Human Factors* 41: 543–552.

Spence, C. 2011. “Cross Modal Correspondences: A Tutorial Review.” *Attention, Perception and Psychophysics* 73: 971–995.

Spence, C., and J. Driver, eds. 2004. *Crossmodal Space and Crossmodal Attention*. USA: Oxford University Press.

Stein, B. E., and M. A. Meredith. 1993. *The Merging of the Senses*. Cambridge, MA: MIT Press.

Szalma, J. L., and P. A. Hancock. 2011. “Noise and Human Performance: A Meta-Analytic Synthesis.” *Psychological Bulletin* 137 (4): 682–707.

Szalma, J. L., and P. A. Hancock. 2012. “What’s All the Noise? Differentiating Dimensions of Acoustic Stress and the Limits to Meta-Analysis.” *Psychological Bulletin*, in press.

Teder-Sälejärvi, W. A., F. D. Russo, J. J. McDonald, and S. A. Hillyard. 2005. “Effects of Spatial Congruity on Audio-Visual Multimodal Integration.” *Journal of Cognitive Neuroscience* 17 (9): 1396–1409.

Tindall-Ford, S., P. Chandler, and J. Sweller. 1997. “When Two Sensory Modes are Better Than One.” *Journal of Experimental Psychology: Applied* 3 (4): 257–287.

Van Erp, J. B. F. 2005. “Presenting Directions with a Vibro-Tactile Torso Display.” *Ergonomics* 48: 302–313.

Van Erp, J. B. F., and H. A. H. C. Van Veen. 2004. “Vibrotactile In-Vehicle Navigation System.” *Transportation Research Part F: Traffic Psychology and Behaviour* 7 (4-5): 247–256.

Van Erp, J. B. F., and H. A. H. C. Van Veen. 2006. “Touch Down: The Effect of Artificial Touch Cues on Orientation in Microgravity.” *NeuroScience Letters* 404: 78–82.

Van Erp, J. B. F., L. Eriksson, B. Levin, O. Carlander, J. E. Veltmanand, and W. K. Vos. 2007. “Tactile Cueing Effects on Performance in Simulated Aerial Combat with High Acceleration.” *Aviation, Space and Environmental Medicine* 78: 1128–1134.

Vitense, H. S., J. A. Jacko, and V. K. Emery. 2003. “Multimodal Feedback: An Assessment of Performance and Mental Workload.” *Ergonomics* 46 (1–3): 68–87.

Warm, J. S., ed. 1984. *Sustained Attention in Human Performance*. New York: Wiley.

Werkhoven, P. J., J. B. F. Van Erp, and Philippi. 2009. “Counting Visual and Tactile Events: The Effect of Attention on Multisensory Integration.” *Attention, Perception and Psychophysics* 71 (8): 1854–1861.

Wolfe, J. M., T. S. Horowitz, and N. Kenner. 2005. “Rare Items Often Missed in Visual Searches.” *Nature* 435: 439–440.